OneCommand™ NIC Teaming and VLAN Manager

Version 2.0.3

for Windows Server 2003

Windows Server 2008

Windows Server 2008 R2

User Manual

P005238-01A Rev. A

One Network.
One Company.
Connect with Emulex.

Copyright © 2003-2010 Emulex. All rights reserved worldwide. No part of this document may be reproduced by any means or translated to any electronic medium without the prior written consent of Emulex.

Information furnished by Emulex is believed to be accurate and reliable. However, no responsibility is assumed by Emulex for its use; or for any infringements of patents or other rights of third parties which may result from its use. No license is granted by implication or otherwise under any patent, copyright or related rights of Emulex.

Emulex, the Emulex logo, AutoPilot Installer, AutoPilot Manager, BlockGuard, Connectivity Continuum, Convergenomics, Emulex Connect, Emulex Secure, EZPilot, FibreSpy, HBAnyware, InSpeed, LightPulse, MultiPulse, OneCommand, OneConnect, One Network. One Company., SBOD, SLI, and VEngine are trademarks of Emulex. All other brand or product names referenced herein are trademarks or registered trademarks of their respective companies or organizations.

Emulex provides this manual "as is" without any warranty of any kind, either expressed or implied, including but not limited to the implied warranties of merchantability or fitness for a particular purpose. Emulex may make improvements and changes to the product described in this manual at any time and without any notice. Emulex assumes no responsibility for its use, nor for any infringements of patents or other rights of third parties that may result. Periodic changes are made to information contained herein; although these changes will be incorporated into new editions of this manual, Emulex disclaims any undertaking to give notice of such changes.

US patent notice is given for one or more of the following: 6226680, 6247060, 6334153, 6389479, 6393487, 6427171, 6427173, 6434620, 6591302, 6658480, 6697868, 6751665, 6757746, 6941386, 6965941, 6687758, 7042898, 7133940, 7124205, 7089326, 6938092, 6996070.

Emulex, 3333 Susan Street
Costa Mesa, CA 92626

Overview 1
Failover (FO) 1
Load Balancing 2
Smart Load Balancing [(SLB)Team Load Balancing] 2
Generic Trunking (Link Aggregation Static Mode) 3
Link Aggregation Control Protocol (LACP) 3
Installation 4
Performing an Unattended NIC Teaming Installation 4
Updating the NIC Teaming Driver 4
Updating the NIC (NDIS Miniport) Driver 4
Uninstalling the NIC Teaming Kit 5
Configuration 6
Starting the Application 6
Device Manager Settings 6
Configuring Teams and VLANs 6
Creating and Configuring a Team 7
Removing an Adapter During Team Creation 8
Primary and Secondary Adapters 8
Configuring a VLAN for an Adapter 9
Team-Adapter Configuration Display 12
Troubleshooting 14

Overview

A VLAN is a network of computers that behave as if they are connected to the same wire even though they may actually be physically located on different segments of a LAN. VLANs are configured through software rather than hardware, which make them extremely flexible. One of the biggest advantages of VLANs is that when a computer is physically moved to another location, it can stay on the same VLAN without any hardware reconfiguration.

In order to increase throughput and bandwidth, and to increase link availability, you can configure multiple network interfaces on one or more CNAs to appear to the network as a single interface. This is referred to as NIC teaming, or multilink trunking.

NIC teaming allows you to group multiple NICs as a single virtual device. Depending on the teaming mode, one or more interfaces can be active. When multiple NICs are combined this way, the group is called a team.

Why Teaming?

NIC teaming has several advantages.

- Increased bandwidth - Two or more network interfaces are combined to share the load, thus increasing bandwidth.
- Load balancing

  Link aggregation enables distribution of processing and communication across multiple links.
- Higher link availability

  Prevents a single link failure from disturbing traffic flow.

Teaming Types

There are different types of teaming:

- Switch independent
  - Failover - If configured for fault tolerance, the system provides only failover.
  - Smart load balancing - If configured for load balancing, failover is included.
- Switch dependent
  - Generic trunking - Link aggregation static mode
  - Link Aggregation Control Protocol (LACP) (802.3ad)

Failover (FO)

Failover is a basic feature of teaming. Though a team may have multiple members, only one member is active at a time. When the active team member disconnects (due to link down, link disabled or any other reason) the failover mechanism selects another team member (which is in linkup state) and traffic continues.

When the primary team member reports a linkup state, traffic fails over to the primary adapter, if the team is configured for auto failback. All the team members use the same MAC address, by default this is the MAC address of the primary team member.

Load Balancing

Smart Load Balancing [(SLB)Team Load Balancing]

Team load balancing provides both load balancing and failover when configured for load balancing, and only failover when configured for fault tolerance.

Team load balancing works with any Ethernet switch and does not require any swtich configuration. The team advertises multiple MAC addresses and one or more IP addresses. The team MAC address is selected by Virtual Team Adapter from the list of load balancing members. When the server receives an address resolution protocol (ARP) request, the software-networking stack always sends an ARP reply with the team MAC address. To begin the load balancing process, the teaming driver modifies this ARP reply by changing the source MAC address to match one of the physical adapters.

Load balancing enables both transmit and receive load balancing based on load balancing function to maintain in-order delivery of frames.

Transmit load balancing is achieved by creating a hashing table using the hashing algorithm based on load distribution type. When a physical adapter (or port) is selected by the Virtual Teaming Adapter to carry all the frames to the destination, the unique MAC address of the physical adapter is included in the frame, and not the team MAC address. This is required to comply with the IEEE 802.3 standard. If two adapters transmit using the same MAC address, then a duplicate MAC address situation would occur that the switch could not handle.

Receive load balancing is achieved through an intermediate driver by sending Gratuitous ARPs on a client by client basis using the unicast address of each client as the destination address of the ARP Request (also known as a Directed ARP). This is considered client load balancing and not traffic load balancing. When the intermediate driver detects a significant load imbalance between the physical adapters in an LB team, it generates Gratuitous ARPs in an effort to redistribute incoming frames. The intermediate driver does not answer ARP Requests; only the software protocol stack provides the required ARP Reply. It is important to understand that receive load balancing is a function of the number of clients that are connecting to the server via the team interface.

Receive load balancing attempts to load balance incoming traffic for client machines across physical ports in the team. It uses a modified gratuitous ARP to advertise a different MAC address for the team IP address in the sender physical and protocol address. This gratuitous ARPs is unicast with the MAC and IP address of a client machine in the target physical and protocol address respectively. This causes the target client to update its ARP cache with a new MAC address map to the team IP address. This has the potential to direct the RX traffic to the different adapter than leaned in from the TX. Gratuitous ARPs are not broadcast because this would cause all clients to send their traffic to the same port. As a result, the benefits achieved through client load balancing would be eliminated, and could cause out of order frame delivery. This receive load balancing scheme works as long as all clients and the teamed server are on the same subnet or broadcast domain.

When the clients and the server are on different subnets, and incoming traffic has to traverse a router, the received traffic destined for the server is not load balanced. The physical adapter that the intermediate driver has selected to carry the IP flow carries all of the traffic. When the router needs to send a frame to the team IP address, it broadcasts an ARP request (if not in the ARP cache). The server software stack generates an ARP reply with the team MAC address, but the intermediate driver modifies the ARP reply and sends it over a particular physical adapter, establishing the flow for that session over the particular physical adapter.

The reason is that ARP is not a routable protocol. It does not have an IP header and therefore is not sent to the router or default gateway. ARP is only a local subnet protocol. In addition, since the Gratuitous ARP is not a broadcast packet, the router does not process it and does not update its own ARP cache.

The only way that the router would process an ARP that is intended for another network device is if it has proxy ARP enabled and the host has no default gateway. This is very rare and not recommended for most applications.

Transmit traffic through a router will be load balanced as transmit load balancing is based on the source and destination IP address and TCP/UDP port number. Since routers do not alter the source and destination IP address, the load balancing algorithm works as intended.

Generic Trunking (Link Aggregation Static Mode)

Link aggregation static is a switch-assisted teaming mode and requires configuring ports at both ends of the link: server and switch ports.

In this mode, the team advertises one MAC address and one IP address when the protocol stack responds to ARP requests. In addition, each physical adapter in the team uses the same team MAC address when transmitting frames. This is possible since the switch at the other end of the link is aware of the teaming mode and handles the use of a single MAC address by every port in the team. The forwarding table in the switch reflects the trunk as a single virtual port.

In this teaming mode, the intermediate driver controls load balancing for outgoing traffic only, while incoming traffic is controlled by the switch firmware and hardware.

Link Aggregation Control Protocol (LACP)

IEEE 802.3ad LACP is similar to link aggregation static mode except that it uses the LACP to negotiate the ports that make up the team. The LACP must be enabled at both the server and the switch for the team to operate. If LACP is not available at both ends of the link, 802.3ad provides a manual aggregation that only requires both ends of the link to be in a link up state.

Because manual aggregation provides for the activation of a member link without performing the LACP message exchanges, it should not be considered as reliable and robust as an LACP negotiated link. LACP automatically determines which member links can be aggregated and then aggregates them. It provides for the controlled addition and removal of physical links for the link aggregation so that no frames are lost or duplicated. The removal of aggregate link members is provided by the marker protocol that can be optionally enabled for LACP-enabled aggregate links.

The link aggregation group advertises a single MAC address for all the ports in the team. The MAC address of the team/aggregator can be the MAC addresses of one of the NICs in the group. The LACP and marker protocols use a multicast destination address.

The link aggregation control function determines which links may be aggregated. It then binds the ports to an aggregator function in the system and monitors conditions to determine if a change in the aggregation group is required. Link aggregation combines the individual capacity of multiple links to form a high performance virtual link. The failure or replacement of a link in an LACP trunk does not cause loss of connectivity. The traffic is simply failed over to the remaining links in the trunk.

Installation

The OneCommand NIC Teaming and Multiple VLAN Manager is installed via a standalone package named `elxdrvr-nic-teaming-<version>.exe`. The driver kit installer is an executable file that self-extracts and copies the following software onto your system:

- NIC Teaming and VLAN driver
- NIC Teaming and VLAN Manager utilities

You may want to unpack all the NIC Teaming software at once, instead of repeatedly installing the kit on various systems. To do this, run the kit and select the **Unpack All Drivers** checkbox on the Installation Options screen.

TCP/IP offload engine (TOE) technology is enabled for Windows Server 2008 R2 systems.

For Windows Server 2003 and all Windows Server 2008 operating systems, you must disable TOE/RSS by applying Microsoft's KB948496.

The NIC Teaming driver supports the following operating systems:

- Windows Server 2003 x86
- Windows Server 2003 x64
- Windows Server 2008 x86 SP2
- Windows Server 2008 x64 SP2
- Windows Server 2008 x64 R2
- Windows Server 2008 R2 with Hyper-V

Performing an Unattended NIC Teaming Installation

An unattended driver installation, sometimes referred to as a quiet or silent installation, requires no user input. This is useful for performing an installation remotely from a command script, or when you want to make sure a custom configuration will not be changed by a user during installation.

To perform an unattended installation, run the driver kit installer with the following command line:

```
elxdrvr-nic-teaming-<version>.exe /q2
```

Updating the NIC Teaming Driver

To update the NIC Teaming driver and preserve the existing configuration (including the IP address), you must first run `elx_octeamupdate.exe` after you install the newer version.

Note: The IP address is not preserved when you update the NIC Teaming driver.

Updating the NIC (NDIS Miniport) Driver

Note: The NIC protocol driver may be updated while NIC teams or VLANs are configured; however, doing so may temporarily disrupt the operation of any teams or VLANs.

To update the NIC (NDIS Miniport) driver, follow these steps

1. Uninstall the NIC Teaming GUI and the NIC Teaming driver by using the 'elx_octeamuninstall.exe' utility, provided with the kit.
2. Update the NDIS driver. See the Emulex Drivers for Windows User Manual for details.
3. Install the NIC Teaming kit (`elxdrvr-nic-teaming-<version.exe`).

Uninstalling the NIC Teaming Kit

Uninstall the NIC Teaming package via the Add/Remove Programs Control Panel (on Windows Server 2003) or the Programs and Features Control Panel (on all Windows Server 2008 operating systems).

Note: Uninstalling the NIC Teaming package will NOT uninstall or delete the current team configuration from your system. To remove any configured teams, you must first run the elx_octeamuninstall.exe utility prior to uninstalling the NIC Teaming kit.

Configuration

A team of adapters functions as a single virtual network interface and appears the same as a non-teamed adapter to other network devices.

A protocol address such as an IP address is usually assigned to the physical adapter. However, when the OneCommand NIC Teaming and Multiple VLAN Manager is installed, the protocol address is assigned to the Team Adapter and not to the physical adapters that make up the team.

The IPCONFIG /all command shows the IP and MAC addresses of the virtual adapter and not of the individual physical adapters.

Starting the Application

From your desktop, click the **OC NIC Teaming Manager** shortcut to start the application.

Device Manager Settings

Before using the OneCommand NIC Teaming and VLAN Manager for simple team or VLAN over team configurations, make sure that the Network Load Balancing and Emulex OneConnect NIC Teaming and Multiple VLAN Teaming checkboxes are cleared on the General tab in the Local Area Connection Properties window.

Configuring Teams and VLANs

Configure teams and VLANs using the NIC Teaming and VLAN Manager dialog box.

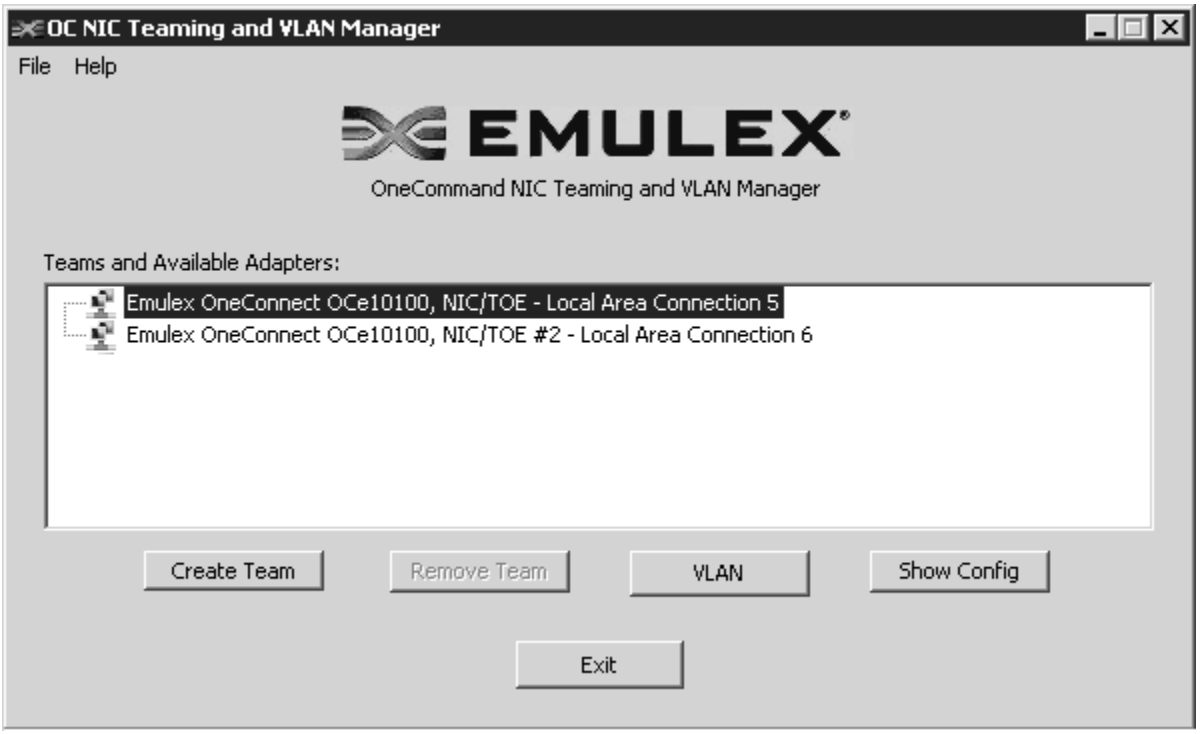

Figure 1: The NIC Teaming and VLAN Manager dialog box

NIC Teaming and VLAN Manager Field Definitions

- Available Adapters - Displays all the physical adapters as well as Team adapters configured, on the system as a tree view. For example:

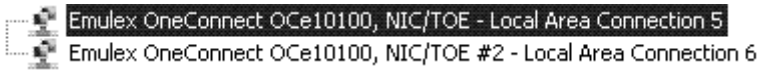
Emulex OneConnect OCe10100, NIC/TOE - Local Area Connection 5
Emulex OneConnect OCe10100, NIC/TOE #2 - Local Area Connection 6

Figure 2: Tree view

NIC Teaming and VLAN Manager buttons

- Create Team - Click to create a new team. The team configuration dialog appears.
- Remove Team - Click to delete a team configured earlier. If no teams have been configured, the button is unavailable. Select the earlier configured team. Click **Remove Team**.
- VLAN - Click to configure VLAN for the selected physical or team adapter:
- Show Config - Click to view adapter configuration. Select the adapter from the list of Available Adapters to view its configuration information.
- Exit - Click to exit the application.

Creating and Configuring a Team

Use the Create Team dialog box to create a team.

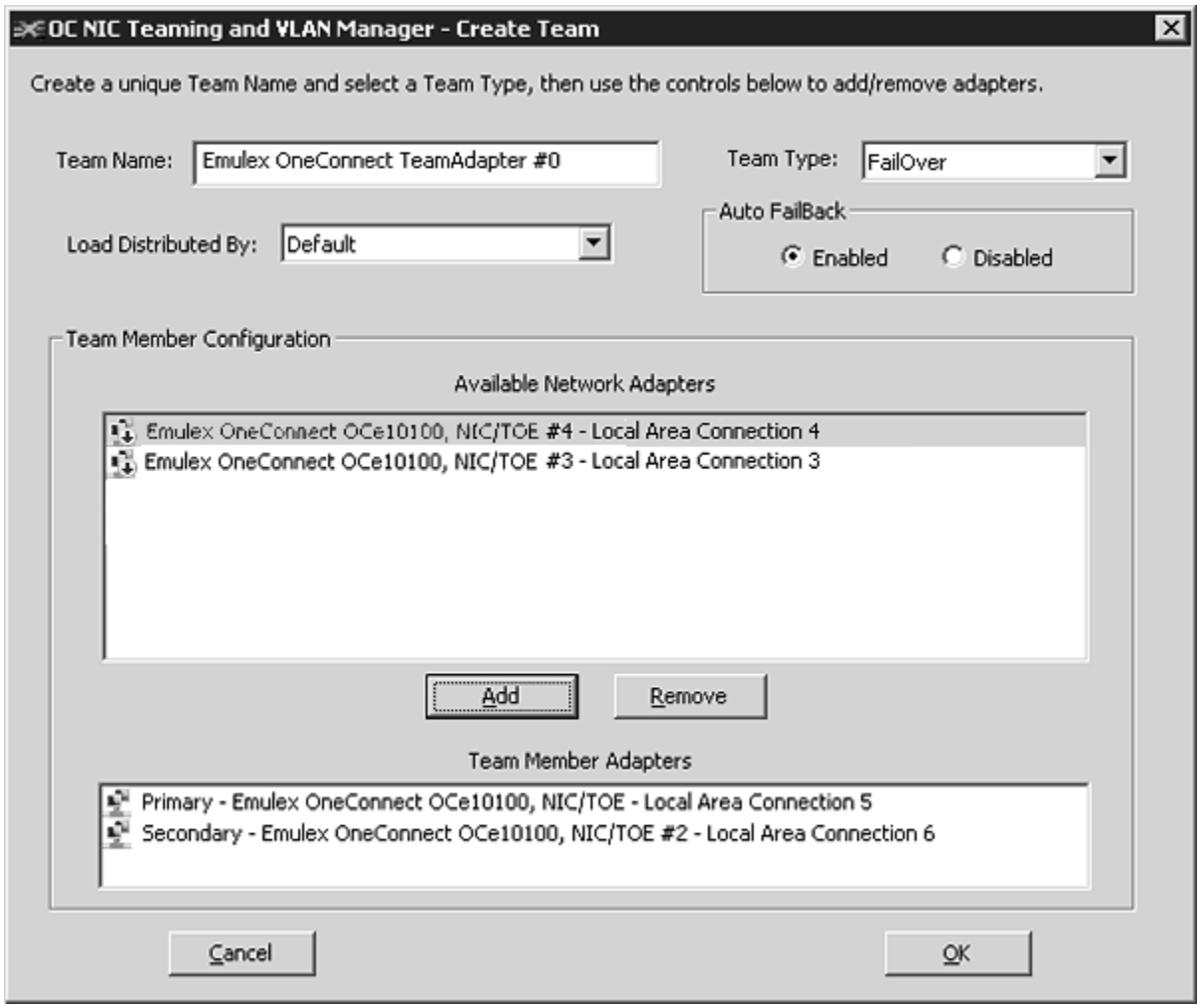
OC NIC Teaming and VLAN Manager - Create Team
Create a unique Team Name and select a Team Type, then use the controls below to add/remove adapters.
Team Name:
Emulex OneConnect TeamAdapter #0
Team Type:
FailOver
Load Distributed By:
Default
Auto FailBack
Enabled
Disabled
Team Member Configuration
Available Network Adapters
Emulex OneConnect OCe10100, NIC/TOE #4 - Local Area Connection 4
Emulex OneConnect OCe10100, NIC/TOE #3 - Local Area Connection 3
Add
Remove
Team Member Adapters
Primary - Emulex OneConnect OCe10100, NIC/TOE - Local Area Connection 5
Secondary - Emulex OneConnect OCe10100, NIC/TOE #2 - Local Area Connection 6
Cancel
OK

Figure 3: Create team dialog box

To create and configure a new team:

1. Enter a new team name. Use the format Emulex OneConnect TeamAadpter #x
   where x is number -0: n based on the last team name number plus 1.
   For example:
   For the first name: Emulex OneConnect Team Aadpter #0
   Subsequent name: Emulex OneConnect Team Aadpter #1
2. Choose a Team Type. Select **Failover > Load balancing > 802.3ad static > LACP** (Default: Failover).
3. If necessary, change the Team Type criteria for Load Distributed By:
   a. Default - Failover. Selects the port with the least traffic load for the session.
   b. Destination MAC Address. Performs an XOR on the destination MAC address to determine which port should carry the load.
   c. Destination IP Address. Performss an XOR on the destination and source IP address to determine which port should carry the load.
4. Enable or disable Auto Failback. The default is enabled.
5. The Available Network Adapters area lists adapters that are not members of any team. To add an adapter to the team, select the adapter from the Available Net work Adapters list and click **Add**. The adapter appears in the Team Member Adapters list.

   Note: Every team must include at least one Emulex adapter.

   The Team Member Adapters area lists adapters that are members of the team.
6. Click **OK** to create a new team.

Removing an Adapter During Team Creation

To remove an adapter during team creation:

1. From the Create Team dialog box, select the adapter you want to remove from Team Member Adapters list.
2. Click **Remove**. The adapter is removed from the Team Member Adapters list.
3. Click **OK**.

Primary and Secondary Adapters

If DHCP is enabled and two ports with IP configuration are added to a failover team, the DHCP assigns the team with the same IP as the active member. This is because the team uses the primary member's MAC address. There is no restriction in using the same IP address of another team member. But to differentiate between a team and non-team member if a team is removed, Emulex recommends using a different IP address.

The first Emulex adapter you add to a team is always the primary adapter. To change the primary adapter, remove the primary adapter from the team and then add the adapter you want to be the primary adapter to the team.

Note: If the adapter you want to be the primary adapter is already a team member, you must remove it and add it to the team.

Configuring a VLAN for an Adapter

All members of the team should have its VLAN disabled at the physical or Windows level. Use the team VLAN configuration, if VLAN is required. Configuring the VLAN at both the physical and team level may cause double tagging.

To configure a VLAN for a physical or team adapter:

1. In the NIC Teaming and VLAN Manager dialog box, see Figure 1, select the physical/team adapter in Available Adapters to which you want to add a VLAN.
2. Click **VLAN**. The Add/Remove VLAN dialog appears. Figure 4 shows the Add/Remove VLAN dialog-box.

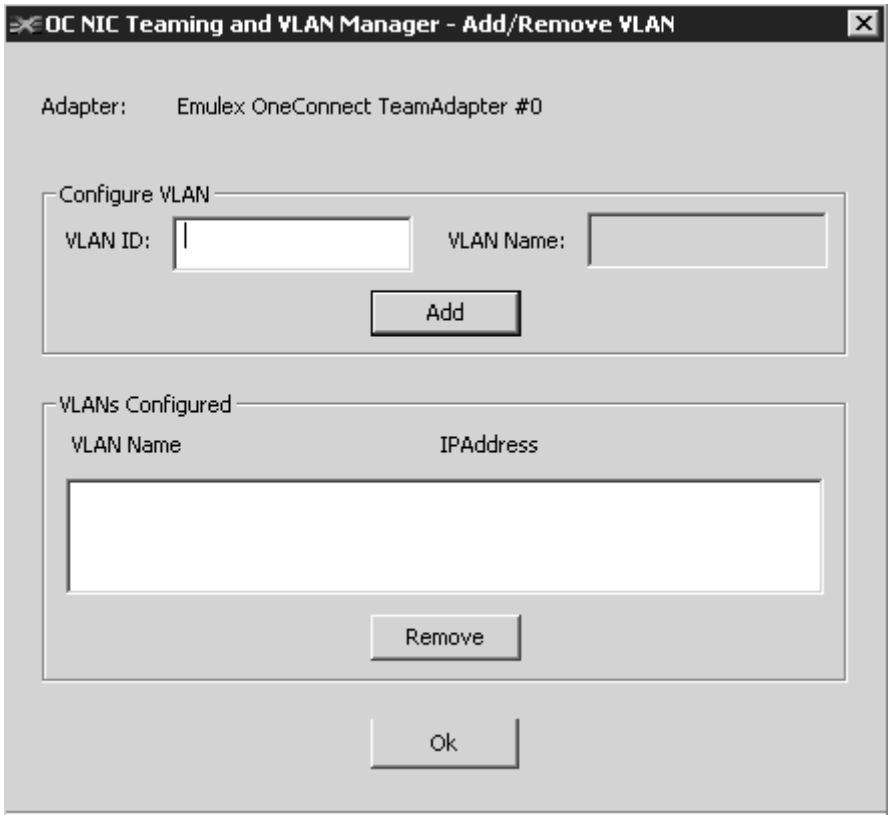

Figure 4: Add/Remove VLAN dialog box

3. Enter a VLAN ID. Enter a VLAN tag (value: 1-4096). The VLAN Name shows the VLAN Name in the format Vlan_<*VLAN ID*>.
4. Click **Add** to add the VLAN to the adapter. You can create multiple VLANs for an adapter. The VLANs Configured list shows the list of all VLANs configured for the adapter.

   Note: To delete a VLAN configured for the adapter, select the VLAN from the VLANs Configured list and click **Remove**.

   Figure 5 shows the newly created VLAN with the list of currently configured VLANs for the adapter.

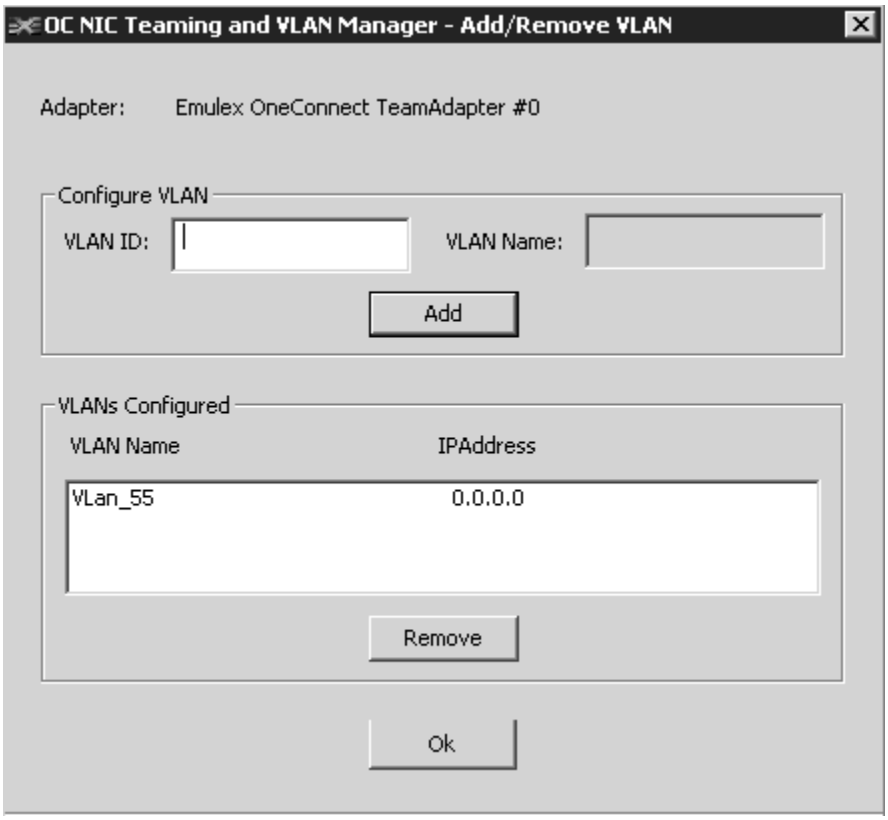

Figure 5: Newly created VLAN screen

5. Click **Ok**. The VLAN is added to the list of configured VLANs.

Note: Once a VLAN is added to a team, the team adapter is marked as down (red x appears in Windows Server 2003 or down arrow in Windows Server 2008).

Figure 6 shows the VLAN in the list.

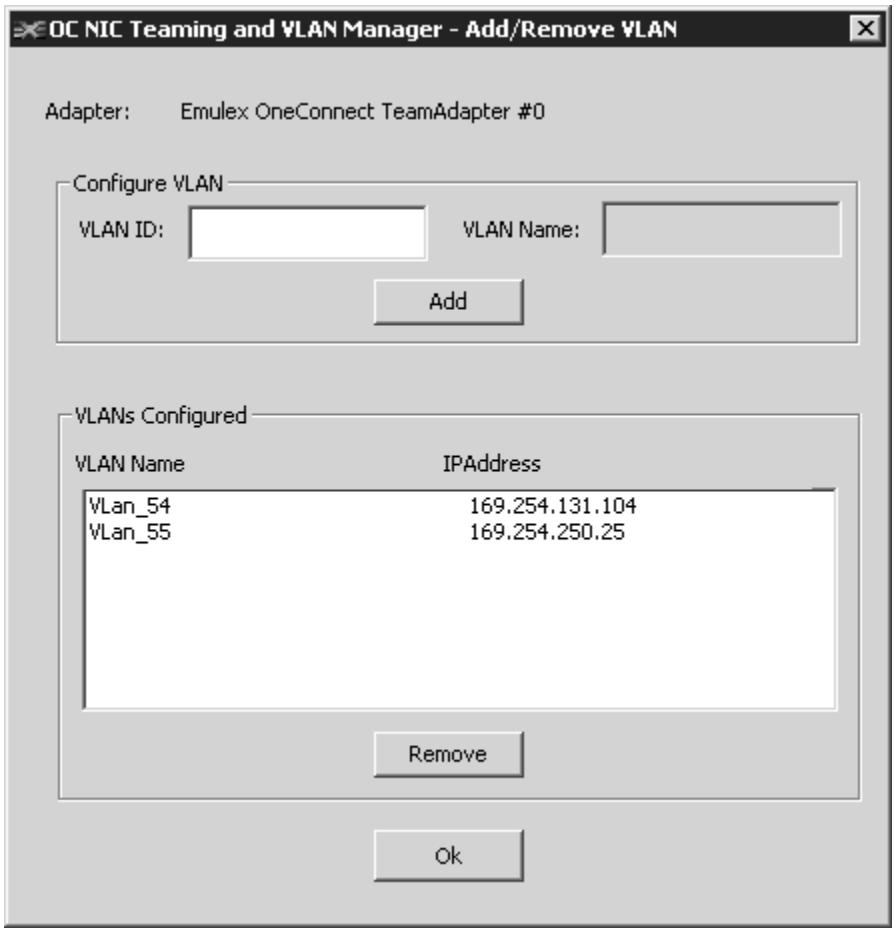

Figure 6: Configured VLANs

Once a VLAN is added to a team, the team adapter is marked as down. A red x appears in Windows Server 2003. A down arrow appear in all Windows Server 2008 operating systems.

Team-Adapter Configuration Display

The Teaming and VLAN Manager Configuration window is read-only and shows the current configuration of the adapter. Figure 7 shows the teaming and VLAN Manager configuration with the primary adapter active.

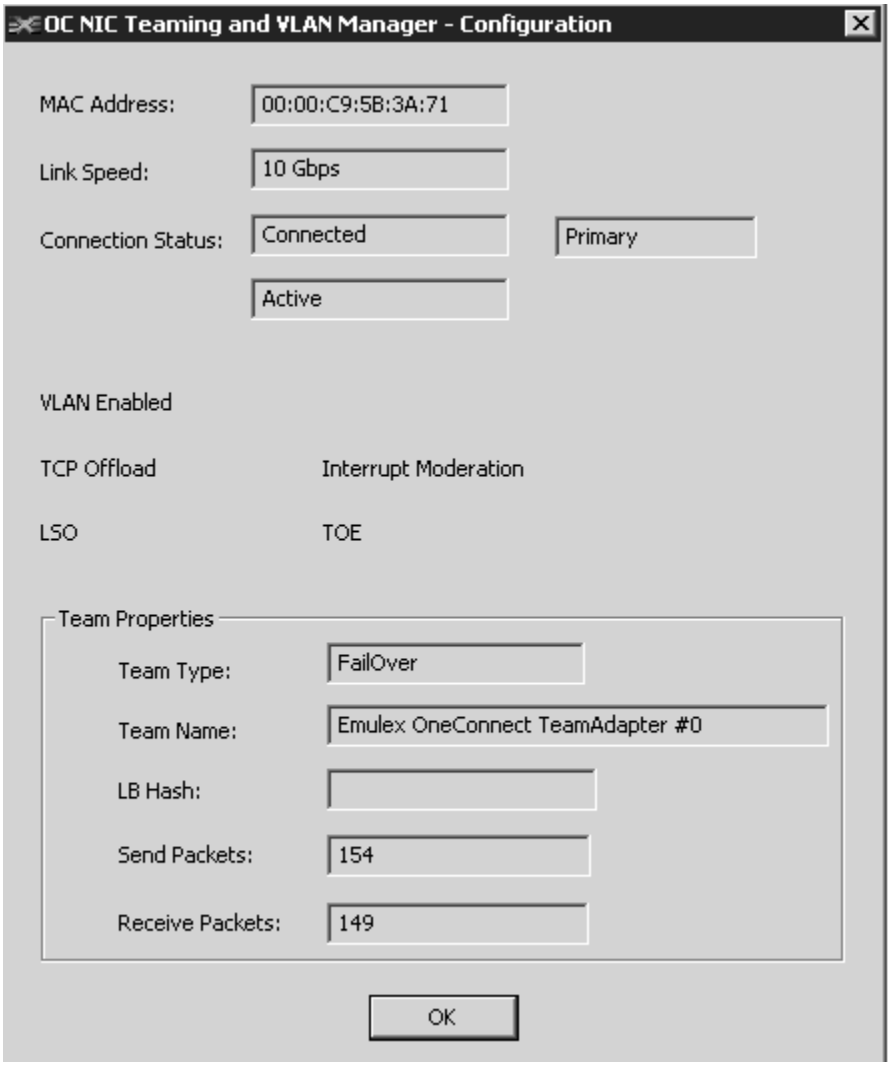

Figure 7: Teaming and VLAN Manager Configuration window

Team Member States

Each team member can be in one of the following states as shown in the Connection Status box:

- **Added:** When a team member successfully binds to the virtual adapter it becomes a member of team.
- **Connecting:** Once the team member is added, the virtual adapter queries it for the state of the physical adapter. At this point, the team member is in the connecting state.
- **Connected:** When the virtual adapter receives the status of the physical adapter, depending on the status returned, it enters either the connected or the disconnected state.

- **Active:** If LACP is set for the team, the LACP handshake starts and LACP adds the link to the team. The adapter's status changes to active, otherwise it remains in standby. In non-LACP mode the connected adapter become active, if you have not set the adapter to standby via the GUI.
- **Disconnected:** When the virtual adapter receives the link-down status indication, (either through LACP or because you disable or remove the link), it is disconnected and it releases the IP address. It is no longer a member of the team.
- **Deleted:** If the adapter is disconnected, you can delete it. If you delete a team member or the whole team, the virtual adapter enters this state. It is no longer a member of the team.

Troubleshooting

There are several circumstances in which your system may operate in an unexpected manner. The Troubleshooting section provides information regarding unusual situations.

NDIS

DESCRIPTION: The team members are incompatible. The Virtual Team driver is NDIS 6.0 and team member (underlying miniport driver) is NDIS 5.0 / NDIS 5.1.

SEVERITY: Information

ACTION: None. NDIS takes care of compatibility issues.

Header Data Split

DESCRIPTION: Incompatible adapters. The team member (underlying miniport driver) supports the Header Data split feature.

SEVERITY: Information

ACTION: None. NDIS 6.1 onwards supports Header Data Split. Windows Server 2008 drivers can handle Header Data Split. If there are any old-NDIS drivers in the network stack, not supporting Header Data Split, then NDIS dynamically turns it OFF.

RSS

DESCRIPTION: RSS (Receive Side Scaling) adversely affects the Virtual Team Miniport Driver (IM driver).

SEVERITY: Information

ACTION: None: In Non 802.3ad mode or 802.3ad static mode, the driver is working as a pass-through driver. So RSS does not affect it. It creates ARP requests and receives ARP replies, but if the OOB fields in the packet are set to 0 (for RSS hash), there is no problem.

In 802.3ad LACP mode creates LACP packets, driver can use the information from OIDs to assign connection on the same CPU that miniport and protocol driver are using. The driver disables RSS by intercepting the OIDS and reports RSS not supported.

Failover

DESCRIPTION: RSS Failover does not work.

SEVERITY: Major

LOG: Teaming members use different hash values.

ACTION: None. A Virtual Teamed adapter sets the RSS capabilities to be similar for all the teaming members using OIDs. The RSS capabilities of the first adapter in a team is used as a base and the RSS capabilities of the subsequent adapters will be compared to it. If different adapters advertise different capabilities, only common capabilities are supported by the Team driver.

Offload features

DESCRIPTION: Adapters cannot be teamed because they support different offloading features.

SEVERITY: Information

ACTION: None. The minimum matching capabilities, among the adapters, becomes the offload capabilities.

SEE ALSO: Table 1

Table 1: Offload Configuration

Offload Capabilities	Available Configuration	Accepted Configuration
1. Checksum Offload	Ipv4, ipv6, TX,RX, TCP.UDP,IP,TCPOPTION, IPOPTION	Based on common factors of TCP/UDP/ IP checksum.
2. Large Send Offload	Ipv4,ipv6 Maxoffloadsize, Minsegmentcount	Based on common factors and minimum of mxoffloadsize and min segment count of all the teamed adapters.
3. TCP Connection offload	Ip4, ipv6 SupportSack TcpConnectionOffloadCapacity	Based on common factors of ipv4/ipv6/ support sack and minimum of TcpConnectionOffloadCapacity.
4. RSS	Base CPU Hash Information Indirection Table Size Indirection Table Offset Secret Key size Secret Key offset	After querying the team adapters, pass the common information to the stack and the pass data from the stack to all the team members (key, key size).

Jumbo frames

DESCRIPTION: Team adapters do not support jumbo frames because they support different offloading features.

SEVERITY: Information

ACTION: Different adapters support different frame sizes. Some support 1514 and 9032 only; some support 2048, 4096 and so forth. The Virtual Team driver considers the minimum packet size reported by OID_GEN_MAXIMUM_FRAME_SIZE and uses it as its MTU. Show Config shows whether the Team adapter supports LSO or not.

VLANs

DESCRIPTION: The adapter supports VLANs, different from the VLAN ID of the virtual team adapter.

SEVERITY: Information

ACTION: If the adapter is configured for VLAN and if you create a VLAN on top of the adapter, data transmission will not work properly.

Multiple VLAN configuration

DESCRIPTION: Must configure multiple VLAN IDs for each virtual team adapter.

SEVERITY: Information

ACTION: Configure an Intermediate Driver (IM) on the top of the virtual team adapter and create VLANs on the virtual team adapter.

Promiscuous mode

DESCRIPTION: Teamed physical adapter is not in promiscuous mode.

SEVERITY: Information

ACTION: None. The teaming/VLAN GUI tries to set the current MAC address for the team member. The driver reads the current MAC address of the adapter. If team MAC address is different from the current MAC address of the adapter, it sets filter type to be Promiscuous mode.

IPv6

DESCRIPTION: Network uses IPv6 - driver incompatible.

SEVERITY: Information

ACTION: None: Windows operating systems, rom VISTA onwards, use only one network stack, so when the driver is bound to the TCP/IP stack, it can receive IPV4/IPV6 packets and it can handle both the packets.

The driver supports IPV6.